MAINTENANCE SCHEDULING OF SPIRAL PIPE MILL MACHINES BASED ON CORRECTIVE MAINTENANCE DATA USING THE RANDOM FOREST MODEL

Nashyh Ulvan Alghany[1], Unit Three Kartini[2]
[12]Department of Electrical Engineering, Universitas Negeri Surabaya
A5 Building Ketintang Campus, Surabaya 60231, Indonesia
[1]nashyh.22087@mhs.unesa.ac.id
[2]unitthree@unesa.ac.id

Abstract – The Spiral pipe mill machine is a vital asset in the steel pipe manufacturing industry that often experiences downtime due to maintenance strategies that are still reactive or corrective maintenance. This approach leads to operational uncertainty and high repair costs. This study aims to create a maintenance schedule for the spiral pipe mill machine using a random forest machine learning algorithm model. This research analyzes 168 historical machine failure data points from January to December 2024 to calculate machine failure intervals. The model was trained using 500 decision trees with an 80% training data and 20% testing data split. Evaluation results show precise model performance with an R-Squared (R^2) value of 0.9916, Mean Absolute Percentage Error (MAPE) of 1.84%, and Mean Absolute Error (MAE) of 0.3993 days. Based on these calculations, an annual maintenance schedule can be compiled to provide accurate maintenance time recommendations to minimize sudden machine failures and increase production efficiency.

Keywords: Maintenance, Maintenance Schedule, Spiral Pipe Mill, Machine Learning, Random Forest.

I. INTRODUCTION

In the era of industrial revolution 4.0, spiral pipe mills play a crucial role in the steel pipe production process. Machines operating with long duty cycles and high loads are susceptible to unexpected breakdowns. When these breakdowns occur, machine downtime is unavoidable. Unplanned downtime can disrupt production stability, disrupt operational continuity, and increase maintenance costs [1]. The global cost of unplanned downtime is US$18.1 billion, which includes labor costs, asset depreciation, and lost sales opportunities[2].

Effective and optimal machine maintenance can be a solution to the problem of production machine downtime. Machine maintenance is a systematic activity of tasks such as monitoring, repair, and replacement designed to maintain or restore the desired function of a machine [3]. Structured and planned maintenance scheduling is the key to reducing costs and increasing machine reliability by preventing unexpected breakdowns[4]. Efficient machine maintenance scheduling can also reduce the risk of failure caused by machine wear [5].

The problem is that many industries in Indonesia still use corrective maintenance strategies. This strategy is less effective because maintenance is carried out when damage has already occurred. This causes high machine downtime, thus disrupting steel pipe production operations [6]. As recorded in the corrective maintenance report data at PT Spindo Tbk. Unit IV, there were 168 breakdowns in one year. This corrective data or historical machine breakdown data contains information on the type of breakdown, the time of the breakdown, and the length of the machine repair. This corrective data contains valuable Time Between Failures (TBF) patterns that can be used to project when similar breakdowns will occur in the future, thus enabling the company to shift from a corrective maintenance strategy to a preventative strategy based on actual data. Scheduling preventive maintenance can help extend the life of your machine by 3-4 times and reduce unexpected machine breakdowns [7].

To process complex and non-linear damage data patterns, an artificial intelligence approach is required, specifically a machine learning algorithm. One algorithm that has proven reliable is Random Forest. This algorithm works using ensemble learning, combining predictions from multiple decision trees to achieve higher and more stable accuracy. [8]. Random forest achieves the highest prediction accuracy in scheduling maintenance on IoT data-driven production systems, far surpassing other models such as XGBoost, MLP, AdaBoost and SVR [9]. Random forest also produces high accuracy in predicting engine failure cycles in the automotive industry, higher than gradient boosting, KNN, and SVM methods [10].

Based on this background, the main problem formulation in this study is how to design a maintenance schedule for a spiral pipe mill machine based on corrective maintenance data using a random forest algorithm model. Therefore, this study aims to create a maintenance schedule for a spiral pipe mill machine by applying the Random Forest algorithm to corrective maintenance data, as well as measuring the accuracy level of the random forest model using standard evaluation metrics such as MAE, MAPE, and R2. In addition, this study also aims to produce a visualization of an annual maintenance schedule that can be used practically as an operational guide to prevent sudden machine failures so as to increase the efficiency of steel pipe production.

II. METHOD

This study uses a quantitative approach with a simulation method to develop a maintenance scheduling model for a spiral pipe mill machine based on corrective maintenance data using a random forest algorithm. The quantitative approach was chosen because this study focuses on processing corrective maintenance numerical data. This study is designed to solve the problem of optimal maintenance scheduling by utilizing historical patterns of machine failure data over a 12-month period. The design of the research flowchart can be seen in the image below.

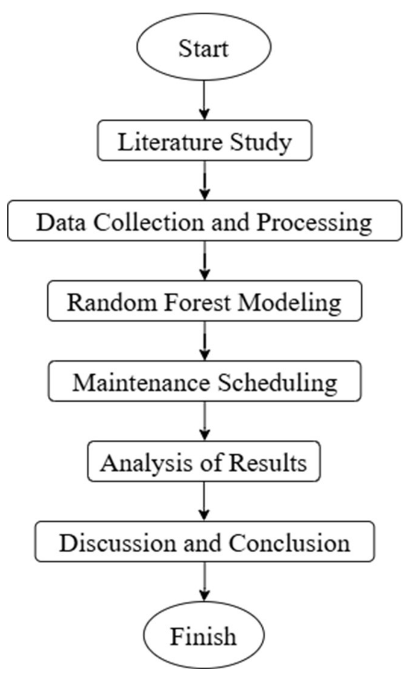

Figure 1. Flowchart of Research Design

The Random Forest modeling process was performed using the Python programming language through the Google Colab platform, utilizing the scikit-learn library (sklearn.ensemble.RandomForestRegressor). The Google Colab platform was chosen because it provides free and easily accessible cloud-based computing and is equipped with the necessary machine learning libraries. Random Forest is a collection of decision tree algorithms combined to achieve better and more accurate predictions. Each decision tree generates a calculation or prediction, and these predictions are then averaged to obtain the final result. The design of the random forest algorithm can be seen in the image below.

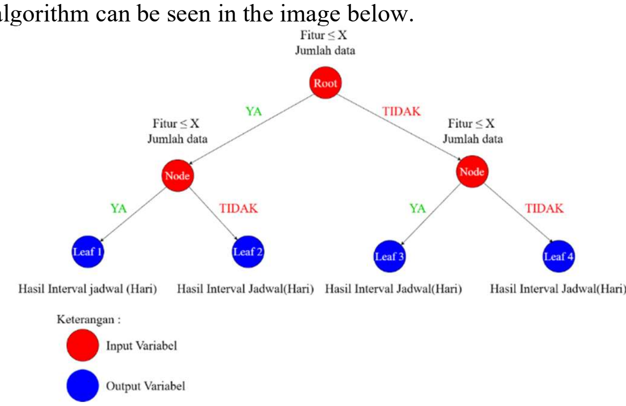

Figure 2. Random Forest Algorithm Design

III. RESULT AND DISCUSSION

Characteristics of Machine Failure Data

The data used in this study is corrective maintenance data for a spiral pipe mill machine collected from January to December 2024. The data comes from maintenance reports that record each machine failure event and the corrective actions taken by the maintenance team. Corrective maintenance data consists of four main variables: the date of the failure, the type of failure or component experiencing problems, the cause of the failure, and the duration of downtime or the length of the machine repair in hours.

To facilitate analysis and the preparation of a structured maintenance schedule, all data is categorized based on the machine component system into 9 (nine) main categories, namely PLC/Main Control, Milling (North/South), Plasma Cutting/Welding, Hydraulic & Pneumatic, Motor & Drive, Sensor & Relay, Mechanical (Bearing/Gearbox), Cable & Connector, and Conveyor System. Categorization is done through text analysis in the damage type and cause columns to identify which system is experiencing problems. Next, the data is input and processed using the python programming language with Google Collaboratory as a development platform for further analysis. The results of the corrective maintenance data analysis are shown in table 1 below.

Table 1 Descriptive Statistics of Data

Parameter	**Value**
Number of Data	168
Period	January 1, 2024 - December 31, 2024
Number of System Categories	9 System Categories
Total Downtime	308.92 hours
Average Downtime	1.84 hours/incident
Minimum Downtime	0.63 hours
Maximum Downtime	2.89 hours

Based on Table 1, 168 failure incidents were recorded throughout 2024, grouped into 9 component system categories.

Cumulatively, these failures resulted in 308.92 hours of total machine downtime. The average downtime per incident was 1.84 hours, indicating that the majority of the disruptions were of a mild to moderate nature, with downtime durations ranging from 0.63 hours to a maximum of 2.89 hours. Although the repair duration per incident is relatively short, the frequency of recurring disruptions can significantly impact overall production efficiency. The aggregated results of component system failure data can be seen in the table below.

Table 2. Component System Aggregation

System Category	Frequence	Total Downtime (Hours)	Average Downtime (Hours)	Damage Interval (Hours)
Cutting Plasma / Las	30	53,49	1.78	12.07
Hidrolik & Pneumatik	25	41.28	1,65	15.00
Motor & Drive	21	42.63	2,03	18.00
Milling System	19	31.55	1,66	19.78
Sensor & Relay	17	29.51	1,74	22.12
Bearing & Gearbox	16	30.01	1.88	24.00
PLC	15	28.58	1.91	25.07
Cables & Connectors	13	26.49	2.04	28.17
Conveyor System	12	25.38	2.11	30.00

Data aggregation is the process of grouping and statistically calculating corrective maintenance data to produce more structured and systematic information. In this study, data aggregation was carried out based on 9 categories of machine component systems to identify damage patterns and downtime characteristics of each system. The aggregation process includes the calculation of 4 (four) important matrices, namely the frequency of damage per component system, total downtime per component system, average downtime per component system, and average damage interval per component system. Then, the distribution of damage frequency in each month can be seen in the figure below.

Figure 3. Frequency of Damage per Month

Based on the image, a clear pattern of fluctuations in the number of machine failures can be observed each month throughout the observation period. In the early to mid-year phase, namely January to July, the number of failures was relatively stable and did not exceed the average value, indicating that the operational condition of the machines was still within controlled limits. During this period, production activities can be said to have run normally, and the maintenance strategy implemented was still able to maintain machine reliability.

However, a different situation was observed from August to November, where the number of breakdowns increased significantly and consistently exceeded the monthly average. This spike in damage indicates an increase in the machine's workload, possibly due to higher production intensity or the accumulation of component wear that is no longer optimally handled by routine maintenance. In addition, the high frequency of damage during this period may also reflect a decline in the effectiveness of reactive maintenance strategies, so that damage tends to occur before maintenance actions are taken.

Based on this pattern, it can be concluded that a more adaptive and planned maintenance approach needs to be implemented, especially during the period from August to November. Strategies such as increasing the frequency of inspections, adjusting preventive maintenance schedules, and strengthening spare parts availability are important to anticipate an increase in the risk of damage, reduce machine downtime, and maintain optimal production process continuity.

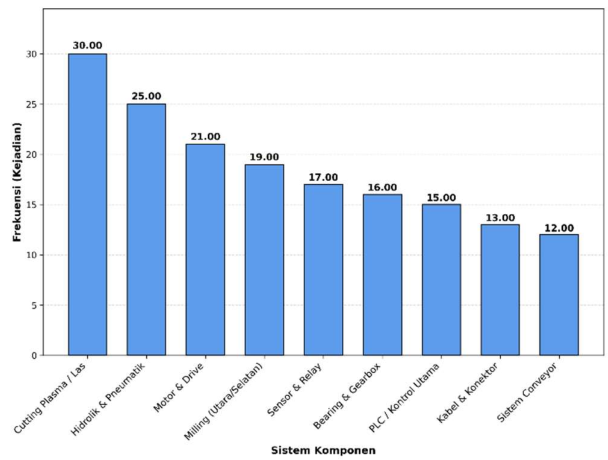

Figure 4. Failure Frequency of each Component System

Figure 4 shows that the plasma/laser cutting system is the component with the highest failure frequency, with a total of 30 failures, followed by the hydraulic and pneumatic systems, which experienced 25 failures. The high frequency of failures in these two systems indicates that these components are operating under heavy operational loads, have high wear rates, or are affected by suboptimal operating conditions. This shows that the plasma/laser cutting system and the hydraulic and pneumatic systems have a relatively high level of criticality and require special attention in maintenance strategies. Furthermore, to understand the impact of these failures on the continuity of the production process, a further analysis was conducted on the downtime caused by each component system. This analysis is presented in the form of a total downtime graph in Figure 5, which provides an overview of the extent to which each system's failure affects the overall machine downtime.

Similar to the frequency of occurrence, the plasma cutting/welding system is identified as the component with the highest total downtime, indicating that failures in this system not only occur frequently but also result in substantial production time losses. This highlights the critical role of the plasma cutting/welding system within the overall production process, where any disruption can significantly affect operational continuity. The motor & drive system ranks second in terms of total downtime. Although its failure frequency is lower than that of the hydraulic & pneumatic system, the longer downtime per incident causes its cumulative downtime to exceed that of the hydraulic & pneumatic system. This suggests that failures in the motor & drive system tend to be more complex and time-consuming to resolve.

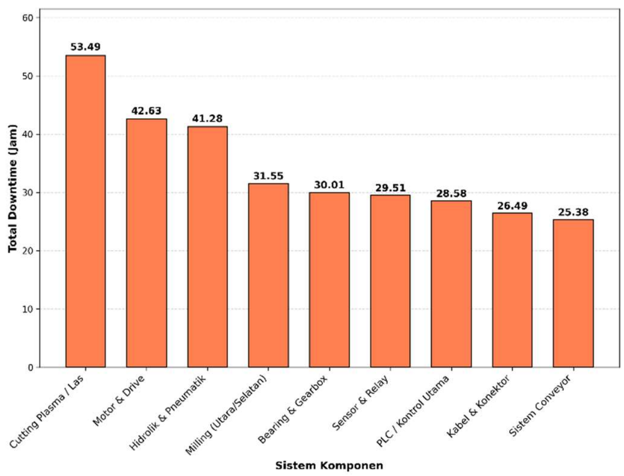

Figure 5. Total Downtime of each Component System

These findings demonstrate that system criticality cannot be assessed solely based on failure frequency, as repair duration and downtime impact play an equally important role in determining the overall effect on production performance. Systems that combine moderate failure frequency with long repair times can contribute more significantly to total production losses than systems that fail more often but are repaired quickly. Therefore, in addition to analyzing total downtime, this study also evaluates the average repair duration or average downtime for each system component to better understand the level of repair complexity and maintenance effort required. The results of this analysis are presented in Figure 6, which provides further insight into the relative difficulty and impact of maintaining each system component.

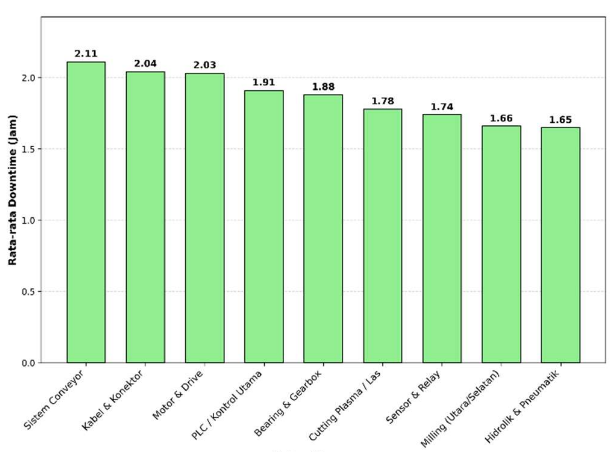

Figure 6. Average Downtime of each Component System

Figure 6 illustrates the average downtime for each component system, which in a maintenance context can be interpreted as the average repair time or Mean Time to Repair (MTTR). The analysis revealed that the conveyor system had the highest MTTR, at approximately 2.5 hours per failure event. This finding is interesting because, although the conveyor system is recorded as the component with the lowest failure frequency, the impact when a failure occurs is relatively significant due to the lengthy repair time required. This condition indicates that damage to the conveyor system is more complex, both in terms of its working mechanism, repair difficulty, and resource requirements, such as technical personnel and spare parts.

Therefore, the conveyor system can be categorized as a critical component that requires special attention in maintenance planning, even though it does not experience frequent failures. Furthermore, to obtain a more comprehensive picture of system reliability, a failure interval analysis was performed, which aims to determine the average time interval between failures in each component system, measured in days. The results of this failure interval analysis are presented in Figure 7 which serves as the basis for determining failure patterns and developing a more effective maintenance schedule.

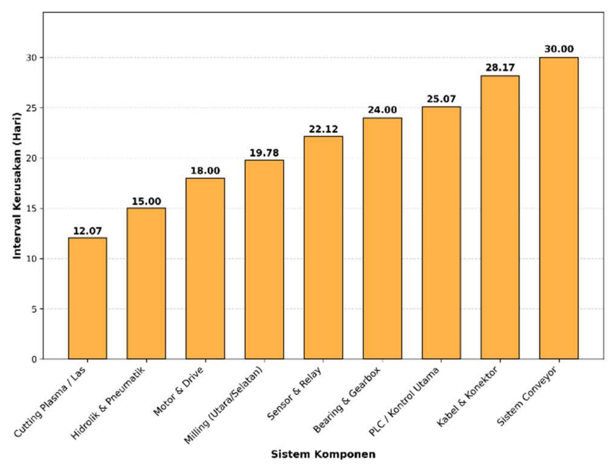

Figure 7. Failure interval of each System Component

Figure 7 presents the average failure interval for each component system in days, illustrating how frequently each component experienced failure during the observation period. Based on the graph, the Plasma Cutting/Welding system had the shortest failure interval, approximately 12 days, indicating that this component experienced failures more frequently than the other systems. This indicates that the Plasma Cutting/Welding system operates under high operational loads and experiences relatively rapid wear and tear, requiring more routine and strictly scheduled maintenance to prevent failures that could disrupt the production process.

Conversely, the conveyor system had the longest failure interval, approximately 30 days, indicating that this system is relatively reliable and experiences fewer failures. However, despite its low failure frequency, the conveyor system still requires special attention because, as previously discussed, repair times tend to be long when failures do occur. Meanwhile, other component systems such as hydraulics & pneumatics, motors & drives, milling, sensors & relays, bearings & gearboxes, PLCs/main controls, and cables & connectors are at intermediate failure intervals, indicating varying levels of reliability and requiring maintenance strategies tailored to the characteristics of each system. Overall, this failure interval information serves as an important basis for determining maintenance priorities, so that maintenance resources can be allocated more effectively and the risk of machine downtime can be minimized.

Random Forest Model Configuration and Evaluation

The Random Forest model was built using 500 decision trees (n_estimators=500) with a training data split of 80% and a test data split of 20%. The selection of 500 trees was determined through an iterative tuning process, which showed that 500 estimators provided the optimal balance between computational efficiency and minimizing the Mean Absolute Error (MAE) without causing the model to overfit. Model performance evaluation was conducted to measure how close the predicted failure intervals were to the actual field data. The test results showed highly precise model performance. Based on statistical evaluation metrics, the model produced a Coefficient of Determination (R^2) value of 0.9916. This value indicates that 99.16% of the failure data variability can be well explained by the model, while the remainder is influenced by other external factors. In addition, the average absolute prediction error MAE (Mean Absolute Error) was recorded at 0.3993 days (less than 10 operational hours), and the average percentage error MAPE (Mean Absolute Percentage Error) was only 1.84%. This low error value proves that the random forest algorithm is able to predict machine failure times with high accuracy, making it suitable for use as a basic reference in preparing maintenance schedules. The results of the random forest model evaluation matrix visualization graph can be seen in the image below.

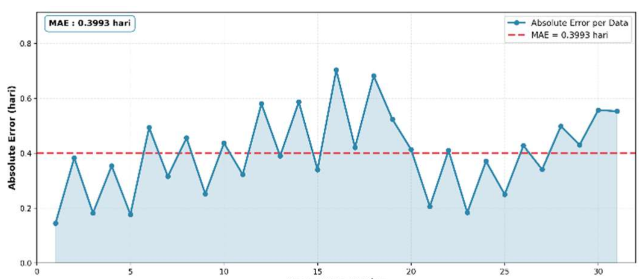

Figure 8. MAE Evaluation Graph

The graph above shows the distribution of absolute errors for each test data set. The X-axis of the graph indicates the number of test data sets, while the Y-axis shows the MAE values for the training data. The red dotted line indicates the mean error (MAE) of 0.3993 days. This figure indicates that the average error in the model's calculations relative to actual field conditions is only around 9.6 hours. In the context of industrial machine maintenance, a time deviation of less than half a day is still very tolerable and allows the maintenance team to prepare for maintenance in a timely manner. The MAPE evaluation graph can be seen in the figure below.

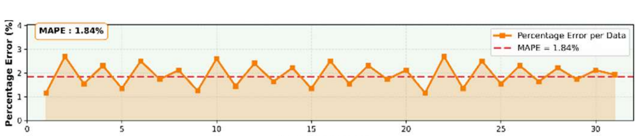

Figure 9. MAPE Evaluation Chart

This graph illustrates the percentage error for each individual test data point, providing a clear overview of how the model's prediction accuracy varies across the dataset. The X-axis represents the order or index of the test data points, while the Y-axis shows the corresponding percentage error measured using the Mean Absolute Percentage Error (MAPE) metric, which reflects the deviation between predicted and actual values. The dashed horizontal line indicates the average MAPE value of 1.84%, serving as a reference for the overall prediction performance of the model.

Referring to the forecast evaluation criteria proposed by Lewis (1982), a MAPE value below 10% is considered highly accurate. Therefore, the average MAPE of 1.84% clearly indicates that the developed model achieves a relatively high level of prediction accuracy. The consistently low error values across most test points indicate that the model is not only accurate on average but also stable in its prediction performance. This low MAPE value further confirms the model's ability to forecast a wide range of failure time intervals, including short-term and long-term intervals, with minimal deviation from the true value. Therefore, the prediction results generated by the model can be considered reliable and robust, making it very suitable to support maintenance planning and decision-making processes, especially in developing proactive and preventive maintenance strategies.

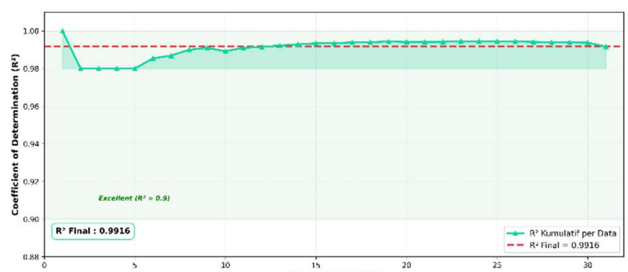

Figure 10. R^2 Evaluation Graph

The graph shows the model's ability to read and analyze variations in the given test data. The X-axis represents the number of test data points, while the Y-axis shows the coefficient of determination (R^2) value generated by the model. The R^2 value is used to illustrate how well the model is able to explain variations in the actual data, where the closer the value is to 1, the better the model's ability to predict. Based on the displayed graph, it can be seen that the R^2 value (indicated by the green line) is in the high range and relatively stable, ending at a final value of 0.9916. The stability of the R^2 value above 0.9 indicates that the model has excellent generalization ability or is robust. This means that the model not only performs well on certain data points but is able to maintain consistent performance across all test data. Although a slight decrease in the R^2 value was observed in some initial data points, this condition is still considered normal in the learning process of a machine learning model and does not significantly affect overall performance. Thus, these results indicate that the model has a high level of reliability in predicting data. Furthermore, to see the agreement between the actual and predicted values in more detail, a comparison of the two is presented in the following graph.

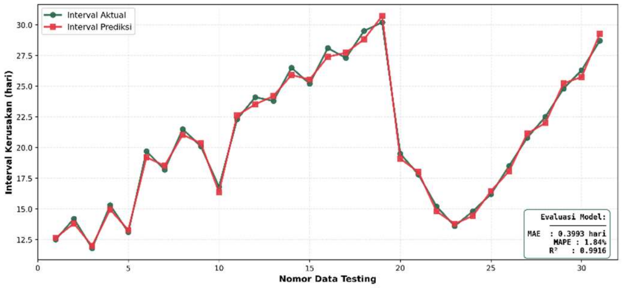

Figure 11. Comparison Chart of Actual and Predicted Values

The figure shows a comparison between the actual failure interval and the predicted failure interval generated by the Random Forest model on the test data. The horizontal axis indicates the test data number, while the vertical axis represents the failure interval in days. In general, it can be seen that the prediction curve has a pattern very close to the actual value curve, which indicates the model's ability to learn the historical pattern of the system's failure interval. At most test points, the difference between the predicted and actual values is relatively small, both at short and long failure intervals. This indicates that the model is able to provide accurate and consistent estimates of the failure interval variation. The model's performance is supported by quantitative evaluation results, with a Mean Absolute Error (MAE) value of 0.3993 days, a Mean Absolute Percentage Error (MAPE) of 1.84%, and a coefficient of determination (R^2) of 0.9916, indicating a very high level of prediction accuracy. With a good level of agreement between the actual and predicted values, the Random Forest model is considered reliable for use as a basis for predictive maintenance schedule planning for spiral pipe mill machines. Furthermore, the combination of a high (R^2) value (0.9916) and significantly low error rates indicated by a MAPE of 1.84% and an MAE of 0.3993 days demonstrates that the implemented Random Forest model achieves superior prediction accuracy and stability compared to similar predictive maintenance models in the manufacturing sector. This robust performance reinforces the algorithm's capability to outperform alternative methods, such as XGBoost or SVM, which have been frequently utilized in previous studies.

Implementation of Machine Maintenance Scheduling

The ultimate goal of this research is to transform the prediction results obtained from the model into a practical maintenance schedule that can be implemented directly in the field. These predictions are used to estimate the time interval for failures in each machine system and component, which then serves as the basis for developing a maintenance schedule. The maintenance schedule is planned to be implemented before the predicted failure occurs, thus minimizing the potential for sudden system failure. This schedule aims to reduce the risk of unplanned downtime, which has historically been a major obstacle to smooth production processes.

Through the implementation of this structured, predictive maintenance schedule, the company is expected to shift its maintenance strategy from a reactive or corrective one, which involves repairs after damage occurs, to a more proactive and preventative one. This change in maintenance approach not only improves system reliability and machine availability but also positively impacts operational efficiency, production continuity, and long-term maintenance cost control. With a planned maintenance schedule, maintenance activities can be carried out in a more systematic and scheduled manner, thus supporting a more stable and sustainable steel pipe production process. The results of the calculation of the time interval for damage to each component, along with the determination of the day for carrying out maintenance, are presented in detail in the following table as a reference for carrying out maintenance in the field.

Table 3. Failure Interval Per System Component

System Category	Damage Interval (days)	Maintenance Schedulu (days)
Cutting Plasma / Las	12.8	10
Hidrolik Pneumatik System	14.90	12
Motor & Drive	17.90	14
Milling	19.75	16
Sensor & Relay	22.06	18
Bearing & Gearbox	23.81	19
PLC / Main Control	25.22	20
Cables & Connectors	27.98	22
Conveyor System	30.15	24

Based on the failure interval per component system generated by the Random Forest algorithm model, a maintenance schedule can then be determined for each component system using a safety factor approach of 80%. The application of this safety factor aims to ensure that maintenance activities are carried out before the actual predicted failure time is reached, thereby minimizing the risk of sudden failure and machine downtime.

The maintenance schedule is calculated by multiplying the failure interval by the safety factor value, resulting in a safer and more planned maintenance time. For example, in a plasma cutting/welding component system that has a failure interval of 12.8 days, a maintenance schedule is obtained every 10 days. A similar approach is applied to all other component systems, such as hydraulic and pneumatic systems, motors and drives, milling, sensors and relays, and conveyor systems. The results of these calculations are then summarized in a table of failure intervals and maintenance schedules per component system. Based on the maintenance schedule for each component, an integrated annual maintenance schedule can then be prepared for the entire system on the spiral pipe mill machine, as visualized in Figure 12. The preparation of this annual maintenance schedule is expected to help companies plan maintenance activities more systematically, improve equipment reliability, and support the smoothness and efficiency of the production process.

As seen in the schedule at figure 12, the plasma cutting/welding system has the highest maintenance frequency, 36 times a year. Conversely, conveyor systems, which have the shortest maintenance interval, every 24 days, only require 15 maintenance sessions per year. This difference in frequency reflects the different levels of operational risk for each system. Systems with short intervals, such as plasma cutting/welding, hydraulics, and pneumatics, are categorized as high-risk components, as their failure can cause significant downtime and production disruption. Conversely, systems like conveyor systems are categorized as low-risk components, due to their longer lifespan and infrequent breakdowns. The schedule compiled above serves as the basis for analyzing the risk levels of component systems in the next section.

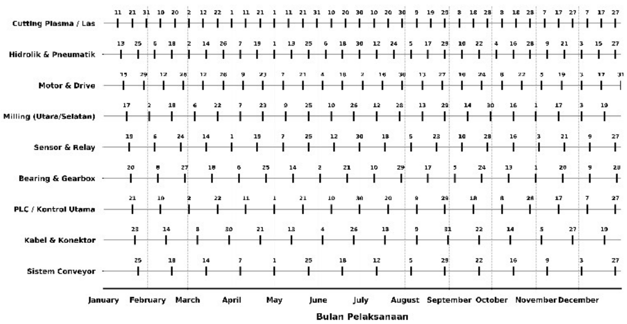

Figure 12. Annual Maintenance Schedule

Component System Risk Level Analysis

After the annual maintenance schedule has been developed, the next step is to conduct a risk analysis for each machine component system. This analysis aims to identify the components with the highest level of criticality to the continued operation of the spiral pipe mill. By understanding the risk level of each system, the company can determine maintenance priorities more objectively, allowing for more effective and planned allocation of resources, such as labor, spare part availability, and machine downtime management.

The risk level is determined by calculating a risk score, which is obtained by multiplying the frequency of failures by the duration of downtime incurred by each component system. This approach is used because a component is considered high-risk not only based on the frequency of failures but also on the magnitude of the impact of these failures on production downtime. Therefore, a component with a low failure frequency but causing long downtimes can still be categorized as a critical component.

The results of this risk score analysis provide a more comprehensive picture of the potential operational risks in each machine component system. This information can be used as the basis for developing priority-based maintenance strategies, such as scheduling more frequent inspections, stocking critical spare parts, and implementing more intensive preventive measures on high-risk systems. This approach is expected to reduce the likelihood of sudden failures, reduce total downtime, and improve overall production process reliability and continuity. A table showing the risk score calculation is shown below.

Table 4 Risk Score System Components

Componen System	Frequency of Damage	Average Downtime	Risk Score	Risk Category
	(F)	(D)	(F × D)	
Cutting Plasma / Las	30	1.78	53.40	High Risk
Hidraulic & Pneumatik	25	1.65	41.25	High Risk
Motor & Drive	21	2.03	42.63	High Risk
Milling	19	1.66	31.54	Medium Risk
Sensor & Relay	17	1.74	29.58	Medium Risk
Bearing & Gearbox	16	1.88	30.08	Medium Risk
PLC / Main Control	15	1.91	28.65	Medium Risk
Cables & Connectors	13	2.04	26.52	Low Risk
Conveyor System	12	2.11	25.32	Low Risk

Based on the comprehensive risk evaluation presented in Table 4, it can be clearly observed that the Plasma Cutting / Welding system possesses the highest Risk Score among all evaluated components, reaching a critical value of 53.40. This is closely followed by the Motor & Drive system with a score of 42.63, and the Hydraulic & Pneumatic system with a score of 41.25. Consequently, these three pivotal systems are systematically categorized within the High-Risk group. This critical classification is primarily driven by their detrimental combination of a high failure frequency and the substantial impact they have on cumulative downtime. In practical terms, this indicates that any unexpected failure or malfunction within these three major systems will inevitably lead to the most severe production disruptions, potentially causing significant operational bottlenecks and financial losses in the manufacturing process.

On the other end of the spectrum, supplementary systems such as the Cable & Connector and the Conveyor System exhibit the lowest Risk Scores in the evaluation matrix, registering values of less than 27. As a result, they are securely categorized as Low-Risk components. Although these low-risk systems still undoubtedly require periodic inspection and maintenance to ensure sustained operational integrity, the frequency of such maintenance interventions can be strategically reduced. This optimization allows the maintenance team to reallocate their focus and resources towards high-risk areas without compromising the overall reliability and performance of the spiral pipe mill machine.

Furthermore, the methodological framework used to classify the risk scores into high, medium, and low categories in this study was determined based on an equal and systematic division of the data into three distinct groups, ranging continuously from the highest recorded risk score down to the lowest. This specific approach was deliberately adopted because, currently, there are no universally accepted international standards or rigid theoretical frameworks that define exact quantitative thresholds for risk classification in this specific machinery context. However, despite the absence of standardized limits, this equal-division method has been widely recognized and practically implemented across various sectors of the manufacturing industry. By utilizing this pragmatic approach, the study ensures a high degree of consistency, objectivity, and reliability in the risk evaluation process, providing a structured and actionable foundation for the subsequent development of the predictive maintenance schedule. The image at figure 13 presents a mapping of the level of damage risk for each component system based on its risk score. This graph uses a traffic light color-coded system to facilitate identification of priority handling, with red bars indicating high risk, yellow indicating medium risk, and green indicating low risk. Based on this graph, several crucial points can be analyzed as follows:

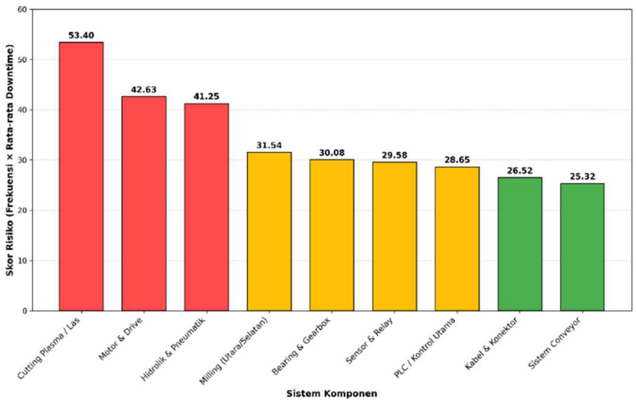

Figure 13. Component System Risk Chart

1. In the High Risk (red) category, three systems fall into the red category with risk scores above 40. Plasma cutting is the most critical (Score 53.40) due to the most frequent failures, up to 30 times per year. Next are Motor & Drive (Score 42.63) and Hydraulics (Score 41.25). It should be noted that motor engines are considered riskier than hydraulics because they require longer average repair times, despite breaking down less frequently. These three systems should be prioritized for spare parts inventory.
2. In the Medium Risk (yellow) category, this group includes four systems: milling, bearings, sensors, and PLCs (Scores 28–32). These systems are at the alert level. The strategy is to conduct regular inspections to prevent minor damage from escalating into high-risk, fatal damage.
3. Low Risk Category (Green): Cables and Conveyors are in the safe (green) category with a score below 28. Failures in these devices have the least impact on production, so they can be handled simply by routine condition monitoring without the need for excessive spare part inventory.

From the analysis of the risk level of each component system above, it can be used as a basis by the company's maintenance team in determining the level of treatment and maintenance schedule for each machine component system, so that it is hoped that the maintenance team can maximize the machine maintenance strategy and minimize the downtime of the spiral pipe mill machine. If the machine downtime can be minimized, then automatically the production effectiveness will be maximized and will reduce maintenance costs to a minimum due to unexpected damage to the spiral pipe mill machine.

IV. CONCLUSION

This research successfully identified the characteristics of Spiral Pipe Mill machine failure data and designed a more effective maintenance scheduling system. Analysis of 168 historical failure data sets revealed fluctuating failure patterns that were difficult to predict using conventional methods. The application of the Random Forest algorithm proved to be an appropriate solution to address these data characteristics, producing highly precise prediction performance. Model evaluation showed a coefficient of determination (R2) of 0.9916, a Mean Absolute Percentage Error (MAPE) of 1.84%, and a Mean Absolute Error (MAE) of 0.3993 days.

This high level of accuracy allows for the development of a valid maintenance schedule that can be used as an operational reference. Implementation of this proposed schedule is expected to shift the company's maintenance strategy from a reactive (corrective) to a proactive one, thereby minimizing unexpected downtime and increasing production efficiency. A limitation of this study is that the predictive maintenance scheduling is solely based on historical corrective maintenance data and does not yet integrate real-time condition-based monitoring utilizing IoT sensors. Therefore, for further development, it is recommended that future research incorporate real-time sensor data alongside more detailed failure variables, such as component types and physical symptoms. Additionally, conducting comparative studies with other algorithms, such as LSTM or SVM, is advised to validate the most optimal predictive model.

ACKNOWLEDGMENT

Thank you to Unit Three Kartini, S.T., M.T., Ph.D., for her guidance, direction, and constructive feedback throughout this research. Her support and motivation were invaluable in completing this research.

REFERENCES

[1] P. Kaźmierczak, K. Żywicki, and P. Rewers, “The Impact of Downtime on the Stability of the Production Schedule,” *Appl. Sci.*, vol. 15, no. 1, Jan. 2025, doi: 10.3390/app15010150.

[2] D. S. Thomas and B. Weiss, “Maintenance Costs and Advanced Maintenance Techniques: Survey and Analysis,” *Int. J. Progn. Heal. Manag.*, vol. 12, no. 1, Apr. 2021, doi: 10.36001/ijphm.2021.v12i1.2883.

[3] M. Jasiulewicz-kaczmarek, “Applied Sciences,” *Early Writings on India*, pp. 124–134, 2024, doi: 10.4324/9781315232140-14.

[4] I. J. O., “Impact of Preventive Maintenance Practices on Productivity,” *Br. J. Manag. Mark. Stud.*, vol. 8, no. 1, pp. 118–137, 2025, doi: 10.52589/bjmms-qrmkxg9r.

[5] D. Ruiz-Hernández, J. M. Pinar-Pérez, and D. Delgado-Gómez, “Multi-machine preventive maintenance scheduling with imperfect interventions: A restless bandit approach,” *Comput. Oper. Res.*, vol. 119, Jul. 2020, doi: 10.1016/j.cor.2020.104927.

[6] G. Oye, E., Ewin, “Transitioning from Corrective to Preventive Maintenance Strategies: Enhancing Equipment Reliability and Efficiency,” no. March, 2025, [Online]. Available: https://www.researchgate.net/publication/389700845

[7] F. Pohan, I. Saputra, and R. Tua, “Scheduling Preventive Maintenance to Determine Maintenance Actions on Screw Press Machine,” *J. Ris. Ilmu Tek.*, vol. 1, no. 1, pp. 1–12, 2023, doi: 10.59976/jurit.v1i1.4.

[8] R. Rosati *et al.*, “From knowledge-based to big data analytic model: a novel IoT and machine learning based decision support system for predictive maintenance in Industry 4.0,” *J. Intell. Manuf.*, vol. 34, no. 1, pp. 107–121, 2023, doi: 10.1007/s10845-022-01960-x.

[9] S. Ayvaz and K. Alpay, “Predictive maintenance system for production lines in manufacturing: A machine learning approach using IoT data in real-time,” *Expert Syst. Appl.*, vol. 173, no. November 2020, p. 114598, 2021, doi: 10.1016/j.eswa.2021.114598.

[10] J. C. O. Ojeda *et al.*, “Application of a Predictive Model to Reduce Unplanned Downtime in Automotive Industry Production Processes: A Sustainability Perspective,” *Sustain.*, vol. 17, no. 9, May 2025, doi: 10.3390/su17093926.